

ОТЗЫВ
официального оппонента на диссертацию
Кашницкого Юрия Савельевича «Методы замкнутых описаний в задаче
классификации данных со сложной структурой», представленной к защите на
соискание учёной степени кандидата технических наук по специальности 05.13.17 –
«Теоретические основы информатики»

1. Актуальность темы диссертации.

Современное состояние информационного общества характеризуется лавинообразным нарастанием объёмов данных, среди которых данные со сложной структурой занимают всё большее место. Это требует создания новых методов обработки таких данных, позволяющих выполнять более глубокий их анализ и, что очень важно, обеспечить интерпретируемость полученных результатов.

Данная диссертация относится к направлению исследований, известному как «Анализ формальных понятий» (АФП). Предметом исследования здесь являются множества данных, связанные отношением «объект - атрибут», единственное ограничение для которых состоит в их упорядоченности отношением частичного порядка. Другой особенностью Анализа формальных понятий является математическая строгость предлагаемых решений и их универсальность, поскольку основой данного направления является математическая теория решёток. Формальный контекст - центральное понятие АФП, определён на произвольных множествах, поэтому может применяться к данным любой природы. Такое преимущество АФП в виде общности организации данных и методов имеет обратную сторону – необходимость адаптации методов АФП к конкретным приложениям, что часто требует специальных исследований.

Тема оппонируемой диссертации и поставленные в соответствии с темой задачи как раз соответствуют данной парадигме Анализа формальных понятий: для решения задач классификации данных со сложной структурой в диссертации формулируется и доказывается ряд утверждений из теории АФП, конкретно - из теории узорных структур. Таким образом, актуальность данной работы обусловлена потребностями практики эксплуатации систем, работающих с данными со сложной структурой, и необходимостью развития элементов теории узорных структур АФП.

2. Содержание диссертации.

Диссертация состоит из введения, четырёх разделов, заключения, включает список литературы и приложение.

Во введении обосновывается актуальность темы исследования, формулируются задачи исследования, описываются применяемые для решения методы, излагаются основные результаты работы, даётся оценка их новизны, научной и практической ценности, приводятся сведения об апробации и внедрении результатов работы, а также обзор содержания работы по разделам.

Первый раздел носит вводный характер и содержит необходимые сведения из теории решёток и АФП. В нём также рассматривается задача классификации и представлен обзор методов классификации, находящихся в поле зрения исследования автора. В конце раздела представлены оригинальные формулировки ДСМ-метода и метода деревьев решений в терминологии Анализа формальных понятий.

В целом, материал первого раздела имеет объем, необходимый для представления содержания работы в последующих разделах, изложен математически корректно, с необходимыми ссылками на литературу.

Второй раздел работы посвящён узорным структурам. Обосновывается необходимость применения узорных структур для решения задач поставленных в исследовании. Вводится ряд определений и доказываются утверждения, служащие далее математической основой алгоритмов, разрабатываемых и исследуемых в последующих разделах. Рассматриваются различные типы узорных структур: интервальные и графовые узорные структуры.

Далее в разделе формулируется задача классификации для узорных структур. Для этого конструктивно используются определения из первого раздела, что обеспечивает связность изложения. Важным результатом данного раздела является установление связи между решётками понятий, строящимися на узорных структурах, и строящимися на формальных контекстах.

Третий раздел диссертации содержит алгоритмы классификации данных, разработанные с применением результатов второго раздела. Предложено два варианта алгоритмов, соответствующих задачам классификации данных с категориальными и количественными признаками на основе множества формальных понятий, а также вариант алгоритма классификации данных на основе узорных структур.

Алгоритмы представлены псевдокодом и описаниями, что позволяет оценить все особенности алгоритмов. Алгоритмы, построенные на основе уже известных алгоритмов, представлены с необходимыми ссылками на их начальные варианты.

Приводятся подробно описываемые примеры применения алгоритмов на данных с категориальными и количественными признаками. На примерах демонстрируется практическое применение результатов предыдущего раздела.

Четвёртый раздел работы содержит результаты экспериментальной проверки алгоритмов, представленных в предыдущем разделе.

В разделе описан программный комплекс, применяющийся в вычислительных экспериментах. Комплекс создан при помощи современных технологий программирования и описан множеством классов.

Эксперименты выполнены на различных тестовых данных авторитетного репозитория UCI. Кроме этого, представлены решения ряда задач классификации: задачи прогнозирования оттока клиентов мобильного оператора, задачи предсказания токсичности химических веществ.

Результаты экспериментов описаны достаточно подробно в виде графиков, таблиц, по данным которых сделаны соответствующие выводы об эффективности разработанных алгоритмов.

В целом содержание диссертации позволяет составить полное представление о разработанных в ней методах и результатах их применения. Текст диссертации написан грамотно, математические определения и результаты изложены достаточно строго.

Список литературы достаточно полно отражает современное состояние в выбранной области исследований, включает как классические работы, монографии, так и последние статьи. Выборочный контроль не обнаружил в списке работ, на которых нет ссылок в тексте.

Автореферат диссертации в полной мере отражает её структуру, содержание, положения и выводы.

3. Обоснованность научных положений, выводов и рекомендаций, сформулированных в диссертации.

Автором исчерпывающе описана исходная проблематика в выбранной области исследования и правильно сформулированы задачи данного исследования. Для решения поставленных задач выбраны соответствующие им подходы и методы. Полученные результаты изложены полно и детально проиллюстрированы в работе. Выводы об эффективности полученных результатов не подлежат сомнению.

4. Достоверность и новизну исследования, полученных результатов, сформулированных в диссертации.

Диссертация в целом представляет собой новое исследование в области Анализа формальных понятий, имеющее практическое значение. В работе получены два новых научных результата:

- новый подход к классификации данных со сложной структурой, не имеющих объектно-признакового описания на основе узорных структур;
- специальный вид проекций узорных структур для данных с количественными признаками, обобщающий подход к обучению на основе деревьев решений.

Достоверность полученных результатов подтверждена строгостью используемых для их получения математических доказательств, экспериментальной проверкой результатов численных расчётов и продемонстрированной практической эффективностью программных реализаций. Также достоверность полученных результатов подтверждается их публичным изложением на конференциях и семинарах и публикациями автора в научной печати уровня, соответствующего требованиям ВАК. Публикации автора отражают содержание и результаты диссертации достаточно полно.

5. Значимость для науки и практики полученных автором результатов.

Полученные автором результаты носят междисциплинарный характер. С одной стороны, в работе развита теория Анализа формальных понятий путём доказательства ряда новых утверждений, проясняющих связи между отдельными разделами АФП. С другой стороны, полученные в работе практические результаты могут служить основой для решения задач классификации данных в самых разных областях: медицине, телекоммуникациях, социальной сфере.

6. Замечания по диссертационной работе.

1. Общее замечание по оформлению работы состоит в том, что её материал и стиль изложения больше соответствует диссертациям на соискание степени кандидата физико-математических, а не технических наук. Однако специальность, по которой защищается автор, «Теоретические основы информатики», возможно, требует именно такого стиля изложения.

2. И в работе, и в автореферате создание комплекса программ для классификации данных со сложной структурой на основе решёток замкнутых описаний представляется в разделе «научная новизна», но это скорее практический, а не научный результат.
3. В списке литературы в публикации [МКК15] не указано издание.
4. Рисунки 4.11 и 4.12 требуют дополнительных пояснений: как вычисляются средние мощности посылок и как интерпретируются изображённые на рисунках фигуры.

7. Заключительная оценка работы

Указанные замечания не влияют на общую положительную оценку диссертационной работы. Диссертация Кашницкого Юрия Савельевича на соискание учёной степени кандидата наук является законченной научно-квалификационной работой, в которой создан новый подход к классификации данных со сложной структурой и разработаны алгоритмы реализации новых решений задач классификации. В целом, в работе сделан теоретический и практический вклад в решение проблемы интерпретации данных со сложной структурой.

Считаю, что данная диссертационная работа полностью отвечает требованиям, пунктов 9-14 Положения о порядке присуждения учёных степеней, утвержденного постановлением Правительства Российской Федерации от 24.09.2013 г. №842, а также требованиям, предъявляемым ВАК к кандидатским диссертациям согласно п. 7 «Положения о порядке присуждения учёных степеней», соответствует специальности 05.13.17 – «Теоретические основы информатики», а её автор, Кашницкий Юрий Савельевич заслуживает присуждения ему учёной степени кандидата технических наук.

Официальный оппонент
доктор технических наук, доцент,
профессор кафедры Информационной безопасности.
Адрес электронной почты: okkambo@mail.ru
Федеральное государственное бюджетное образовательное
учреждение высшего образования «Тульский государственный университет»
Адрес: 300012, г. Тула, пр. Ленина, 92
<http://tsu.tula.ru>

Богатырев Михаил Юрьевич,
24.09.2018

Подпись профессора М.Ю. Богатырева подтверждена:

Начальник

Административно-кадрового управления

М. В. Метелищенкова



24.09.18